



# ENTERED

PCT10

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/018,470A

DATE: 02/27/2003 TIME: 08:05:17

Input Set : N:\Crf4\02132003\J018470.raw
Output Set: N:\CRF4\02272003\J018470A.raw

```
1 :110 > APPLICANT: Frazer, Claire M.
         Hickey, Erin
         Peterson, Jeremy
         Tettelin, Herve
         Venter, J. Craig
 ŧ,
         Masignani, Vega
         Galeotti, Cesira
         Mora, Manrosa
 Я
         Ratti, Giulio
10
         Scarselli, Maria
         Scarlato, Vincenzo
11.
1J
         Rappuoli, Rino
         Pizza, Mariagratia
13
14
         Grandi, Guido
15 < 120 >  TITLE OF INVENTION: Neisseria Genomic Research
16 <130 > FILE REFERENCE: CHIR-0319
17 <140> CURRENT APPLICATION NUMBER: US/10/018,470A
18 -141 CURRENT FILING DATE: 2000-03-08
19 < 1500 PRIOR APPLICATION NUMBER: 60/132,068
20 (151) PRIOR FILING DATE: 1999-04-30
21 <150% PRIOR APPLICATION NUMBER: PCT/US99/25373
22 - 151 - PRIOR FILING DATE: 1999-10-08
23 ×1500 PRIOR APPLICATION NUMBER: GB-0004695.3
24 (151) PRIOR FILING DATE: 2000-02-28
25 -150 PRIOR APPLICATION NUMBER: PCT/US/05928
26 +1151 - PRIOR FILING DATE: 2000-03-08
17 - 160 NUMBER OF SEO 1D NOS: 107
28 - 170 - SOFTWARE: PatentIn version 3.1
30 <210: SEQ ID NO: 1
31 -: 211. LENGTH: 2272325
32 - 212 - TYPE: DNA
33 - (213) ORGANISM: Neisseria meningitidis
34 -: 400> SEQUENCE: 1
                                                                                  60
35
         taaaccttat ccacatecaa acqcataacc qtaacccatt caccqttatq qaaatqtcqc
3 ti
         cogacaacca cocagoogaa tgattoataa aatatttgoa catcaggogt ataaagatac
                                                                                  120
                                                                                  180
37
         aaqaacttta teeccaqeqa aeqeqetqeq eetatqeaqt qqqeqaceaq eeteetgeca
                                                                                  240
3 11
         atgeettite egogatatie aggitaaaaea aagacateee eeaaccaata ticatacegit
                                                                                  300
34
         ggaaaacttt ecatateatg cegettgaec geageegaac ceaacaggat teeggaatea
                                                                                  360
         tocacagoog caaatgecag eggeagtteg teatcettea aacacetgee gtaataggea
41
         tgaatottat ocacagaaga ocacggttoa aatoogtgoo actootoaaa caacgootga
                                                                                  480
4.3
         accaacctgc cyatatgccc ggctttcagc cgtgtaatga aaacagtatt gtccacaaag
                                                                                  540
4 3
         agggaattea teggteaatt eecegaegee ttegtteeec etgegeegta aacegeatte
```

caaqcatqqt ccaaacqcac tecqatttqc ctcaaatctt caqcctqccq gqctttttqc

600

4.1

### RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/018,470A

DATE: 02/27/2003 TIME: 08:05:17

Input Set : N:\Crf4\02132003\J018470.raw
Output Set: N:\CRF4\02272003\J018470A.raw

						L	3600
94	ggcttacaca	gggttttact	caatatcccg	cctacaaccg	taccaaacgg	LLLacaalac	3660
95 •	ccgaatcgac	atacaaagga	caaaacgatg	aaatacttga	atcttgccgc	aatcaccctt	3720
96	gccgccacat	ttgccgcaca	taccgcctcg	gcagacgaac	tggccggatg	gaaagacaac	
97	accccgcaaa	gcctgcaatc	geteaaagee	cccgtacgca	tegteaacet	ttgggcgact	3780
98 -	tggtgcggcc	cgtgccgaaa	agagatgcct	gccatgtcca	aatggtacaa	agcgcagaaa	3840
99	aaaggcagcg	tcgatatggt	cggcat.cgcg	ctcgacacat	ccgacaatat	eggcaacttc	3900
100	ctcaaacaaa	ctcctgtttc	ctacccgatt	. tggcgttaca	ccggggcgaa	cagccgaaac	3960
101	tttatgaaaa	cctacggaaa	cactgtcggc	gtactgccct	ttaccgtcgt	cgaagcaccg	4020
102	aaatgeggat	acaggcagac	cattaccggg	gaggtaaacg	aaaaaagcct	gaccgacgcc	4080
103	gtcaaactcg	cccattcaaa	atgccgttaa	acgccggatg	cegtetgaag	ccgcttcaga	4140
104	tggcattttt	cttttccacc	cgcctgccgg	, tgcaaactta	tccactatct	aaaaacaggc	4200
105	ggaatcttta	taatcggcac	tgtcttacct	. attgttcaga	cggcatatcc	ctgcggacgc	4260
106	aaccgcccga	aacgatatgo	cgcccttcct	. tacaggacct	. cctatgatcc	gtttcgaaca	4320
107	agtttccaaa	acctatcccg	gcggttttga	agccctgaaa	aacgtcagct	tccaaatcaa	4380
108	caaaggcgaa	atgatattta	tegegggaca	ctccggttcg	, ggcaaatcca	ccatcctcaa	4440
109	actgatttcg	ggcattacca	agccgagcag	gggcaaaato	ctgtttaacg	ggcaggacct	4500
110	eggeaeattg	teegaeaace	aaatcggctt	tatgcgccaa	ı cacatcggca	tcgtgttcca	4560
111	agaccacaaa	atcctctacg	accgcaacgt	cctgcaaaac	gtcatcctgc	cgcttcggat	4620
112	tatoggotat	ccqccqcqca	aagccgaaga	gegtgeeege	atcgccatcg	aaaaagtcgg	4680
113	cctgaaagga	cgagaattgg	acgaticccgt	aaccctctcc	ggcggtgaac	aacaacgcct	4740
114	gtgcatcgcc	cacaccatca	ttcaccagec	cggcctgctg	g attgccgacg	aaccctccgc	4800
115	caacetegae	cacacctaca	cgctcgatat	tatggaattg	, ttcaaaacct	tccacgaagc	1860
116	gggaactacc	gtcatcgttg	ccqcacatqa	cgaaaccct	g atggcggact	acggacaccg	4920
117	catectgege	ctctcgaaag	gacgaetege	atgagcatca	tccactacct	ctcgctgcac	4980
118	gt.cgaat.ccg	cacacaccac	geteaageag	ctectgege	aaccettegg	cacactgctt	5040
119	acceteatga	tactcaccat	. cgcgatgacc	ctaccactat	ttatgcatct	gggcatccaa	5100
120	адсиддсааа	acatattaga	caaactcaac	gagtegeege	c aaatcacaat	ctatatggaa	5160
121	acctccacca	: сасааадсда	cagedatace	gtccgcagc	tgctggcgcg	cgacaaacgg	5220
122	ctcgacaaca	tecaetteat	coocaaagaa	gacggtctg	, aagaattaca	gtccaatctt	5280
123	gaccaaaatc	tgatttccat	gettgaegge	aaccccctg	cggatgtctt	tateqttace	5340
123	cccaacccaa	caaccacqcc	: cocccaaato	r caggeaatét	accyagacat	taccaaactg	5400
125	cctataatca	aatccgcgtc	: tatugataco	gaatgggtg	aaacgctgta	ccaaatcaac	5460
126	gagttcatcc	gcaaaatttt	ataatttett	tecetgacge	tggggatggc	gttcgtcctt	5520
127	gageeaacaca	acaccatcco	r cotgeaaato	ctcagccgca	aagaagaaat	cqaaatcacc	5580
1.28	aaactettaa	acacacacaca;	grouttate	e egeegeecat	tecttiatea	agccatgtgg	5640
129	cararcatic	tttccaccac	cateaactte	ı agaetttac	g gttggctgct	ctctaccata	5700
130	oggageattaa	tegatgeeat	tttcaaacc	r tacqqactta	atateggetg	acaattette	5,71,1)
131.	taggetggg	aactegggeet	antatteaa	e tteateatea	g cgttgggcgt	atteggegeg	5820
132	tacyctygcy	addeegggee	. ggtgttegg:	ttcaaagca	a aaaaataaaa	caccatcaaa	5880
133	antagagtag	. coaccoage	t cagacaga.	e tttcaattt	g ccagtataat	ggcgcatttt	5940
	taggeegeee	. gaaccegeet	actuacetee	r daacaadta	a aagccatgat	tgaaggcgtg	6000
134	rannantaa	adctaccac	. googaceee	g gaasaageas	e atttttcgc	catcatcatt	6060
1 3 5	geadaatgeg	ttannagan	agcagaagge	gacggacae.	geetgattaa	agacquacto	6120
136	Loatoayadt	. Lugaayyuda	i ggedegeete	gegegeedet	tttcggttgc	caccactcca	6180
137	aaaycccaac	: cyyadayida	i ogađetjede Lagaataate	T GCaCaCaCaaa	a ataccateta	agaccattte	6240
138	geggaatggg	cagodaaago	: acaalaalC	g coactagaaa	a atgeegtetg	ссазапедде	6300
139	gtttcagacg	geautititi	. cataccada . attacaaca	ot da adda a	g cogogittit	caataaaaat	6360
$\frac{140}{140}$	tacggcaggc	agelettige	. agaaraatt	a cicaayyaa	e gegeegeege	caccaccacc	6420
141	gtageegatt	. igticggtaa	. ttt zaaatt	_ ggcaatcgcc	gecagegtgt	taccacctes	6480
142	cgcaatcgag	aacgetttge	: LLLyggcaat	_ ggcttegge	a agggettteg	caccyccigo	0400

## RAW SEQUENCE LISTING

DATE: 02/27/2003 TIME: 08:05:17 PATENT APPLICATION: US/10/018,470A

Input Set : N:\Crf4\02132003\J018470.raw Output Set: N:\CRF4\02272003\J018470A.raw

143	gaattggtca	aactcgaaca	cgccgaccgg	cccgttccaa	acgaccgtac	cggcggcttt	6540
144	aagcaaatcg	gcaagcgcgg	cagcggattt	cggaccgatg	tccaaaatca	tctcgtcttc	6600
145	ggcaacgtcg	gcaatgtctt	tcaccacage	ttccgcatcg	gcggcaatgt	ctttcaccac	6660
146	agetteegea	teggeggeaa	aggetttgge	aacgacgaca	tcggtcggca	gcggcacaga	6720
147	accgcctttt	gccgccattt	tegecataat	ttttttggat	tcttccacca	aatcgtgttc	6780
148	cgccaaagat	ttgccgatgg	ctttgccttc	cgccaacagg	aaggtgtttg	cgataccgcc	6840
149	geegaegatg	agttggtcga	ctttgtccgc	cagcgattcg	aggatggtca	gcttggtgga	69110
150	cactttgctg	ccggcaacga	tggcaaccat	cgggcgcgcg	ggctgtttca	aggctttgcc	6960
151	caaagcgtcg	agttcgcccg	ccatcaatac	gccggcgcag	gcaacgggcg	cggcttgggc	7020
152	gacggcttcg	gtcgaggctt	gggcgcggtg	ggcggttccg	aacgcgtcat	tgacgaacac	7080
153	gtcgcacaaa	gaagcgtagg	ctttacccag	ttccaaatcg	tttttcttct	cgcctttgtt	7140
154	gatgcgcacg	ttttgcagca	tgacgacatc	gcccgcgttc	agggcgggtt	tgttttcacg	7200
155	ccagtcgttc	aatactttca	cgtctttgcc	caacaggctg	cccaagtgcg	cggcaacggg	7260
156	ggcgacatcg	tcttcggggt	ggaactcgcc	ttcggtcggg	cggccgagat	gggtcatcac	7320
157	gataacggac	gcaccgttgt	ccacgcagta	tttaatggac	gcgagcgagg	cgcggatacg	7380
158	ggtgtcgtcg	ctgattttgc	cgtctttgaa	cggtacgttc	atatcggcgc	ggatgaggac	7440
159	ggttttgccc	tgcacgtttt	gttcggtcag	ttttaaaaat	gccataatca	gtccttttca	7500
160	atcagtgttt	gcgatacgga	aacaattgat	gccgtctgaa	ggcttcagac	ggcatcycaa	7560
161	cccgatcage	cggatacgcg	ctcgattttc	gegeegaege	tgccgagttt	tttttcaata	7620
162	ttttcataac	cgcgatccaa	gtggtaaatc	tgttcgacca	cggtttcgcc	tegegeegee	7680
163	aaaccggcga	taacgaggct	ggcggacgca	cgcaaatccg	tegeettgae	gactgcgccg	774()
164	gaaagctgtt	ccacaccctg	cacaaatgcc	gtattgccct	cggttgtgat	gttcgccccc	7800
165	atccggttca	actcggggac	gtgcataaag	cggttttcaa	aaatcgtttc	caccacgcgg	7860
166	cagetteect	ccgccacggc	attcaatgcc	ataaactgcg	cctgcatatc	cgtggggaag	7920
167	ccggggtgga	cgaccgtgcg	gatgtccacc	gccttcggac	gctgccgcat	atcgatggcg	7980
168	atccaatcgt	cgcccgcctc	aatcaccgca	cctgcctcaa	ccagtttgtc	caacaccact	8040
169	tocatogttt	teggegegge	attccgcaaa	accaccctgc	caccggttat	cgccaccgcg	8100
170	cacaggaacg	teccegeete	gatccggtcg	gggacgacgc	tgtgttcgca	gccttgcagc	8160
171	tcgtccaccc	cttccacaat	cattgtggac	gtaccgatgc	cgctgatttt	cgcgcccatt	8220
172	ttgaccaggc	attccgccaa	atcgaccact	tcaggctcaa	tggcgcagtt	ttccaaaacc	8780
173	gtcgtacctt	cegecagegt	egeegecate	agcaggtttt	ccgtgccgcc	gacggtaacg	8340
174	acatccatcg	ccacgcgcgt	acctttgagt	ttgcctttgg	ctttgacgta	accgtgttcg	8400
175			egetteeaag				8460
17h	ccgatggcgc	ageegeeegg	caggetgaet	tgagaataga	cgaaacgcgc	cagcgtcggg	8520
177	-		gegeategtt				8580
178	-		ttcaaattcg	-			8640
174	ateceetgaa	geagettitg	cgtggttttc	acatotgoca	gcatagggac	gtttttcagg	8700
180			caaacccgcg				8760
181			gttgagcggg				8820
182			acttgtatag				0888
183			agtacagaac				8940
184			ggcgaggcaa				9000
185	atatttcaat	tetegggaea	acgcataaag	catcaccega	tgaaggttgc	agaggcggaa	9060
186			aaatacggaa				91.20
187			cggaagacaa				9180
188			gagegatteg				9240
189			gacccgataa				9300
190			gacgggcgtt				9360
191	taaaccaaag	aagtcaacac	cgatgcaccc	aatgcgccga	acacgaccga	cagcgaaacg	9420

#### RAW SEQUENCE LISTING

DATE: 02/27/2003 PATENT APPLICATION: US/10/018,470A TIME: 08:05:17

Input Set : N:\Crf4\02132003\J018470.raw Output Set: N:\CRF4\02272003\J018470A.raw

192	gaaatcggga	tatgcaccca	atgeattace	agcattttca	caccgataaa	acccaacacg	948()
193			gaagataaag				9540
194	ategecegea	agcccagaat	tgcgaaaata	ttggaagtca	gcacgataaa	cggatcggtg	чы́()()
195			gatgctgtcc				9660
195	atgaccagca	ccaaaaacag	cggcgtggcg	atttttttgc	cgttttcgac	ggtaaaaaat	9720
197	ttctcgccgt	gaaattccgt	geegaeegga	acgactttct	tgacggtatt	cagcagcctg	9780
198	ctgtttgcca	aatcctcttt	ctcatcgcct	tegggettea	tcatgtgtat	accagtatag	4840
199	agcaggaacg	egecaaacag	atacagaatc	cactcaaact	gctgaaccag	tgccgcgccg	9900
200	acgaaaatca	tgacggtgcg	caataccaat	gcgcccaata	cgccgtacag	cagcacgcgg	996()
201	tgctgaaact	gtggtgcgac	tttgaagtag	ccgaatatca	tcaggaacac	gaaaatattg	10020
202	togactgoca	acgatttttc	caaaatgtag	ccggtaaaga	attccaatac	tttttcttt	10080
203	gcgactgccg	cgccgtagcc	gggattgccg	gcgagttcaa	aatacagcca	gcccgcgaac	10140
204	aggcaggata	cggcaaccca	caageegete	catgccaagg	cttctttgac	gccgacttta	10200
205	tggctgccgt	ttttcttcag	cgaaaacata	tocaaggcaa	tcatgaccag	cactgccgca	10260
206	aaaaaaacgc	cgtaaaacaa	cggcgacccg	atgccgggat	attctgtcat	ggttcaatct	10320
207	cctgatttga	aatgtaattg	tgttaccagc	tgatataaaa	catcgctttt	gccaaaaaga	10380
208	caatcagcag	catatgggta	aagacgacgg	cgtgtatgta	tttcgaccaa	ccgaccgtca	10440
209	gtgtggaacg	cgccattttg	acgacggcga	tggcgaagtg	cgccaatacg	ctgaacgcca	10500
210	acaggatttt	cagcgtcagc	atcgtaccga	aggaagtggc	aaacggttcg	cccaatatag	10560
211	aaagatagcg	gtttgccgcc	atcacgatgc	cgctggcgaa	cagcagtccg	accacaaacg	10620
212			taagacattg				10680
213	acacccgtcc	cgtatgcagg	acggacaaaa	ccagcacttc	aaaaaacacg	ccgccgacaa	10740
214			tgaacgatgt				10800
215	ccaaacygaa	aactcggata	cggattgtat	cactatoged	cccgatatcc	gcataccgct	10860
216			totogogoco				10920
217			ccccccccc				10980
218	cgccacgcat	ccgccgattc	gacttgaacc	geogeacceg	atgccaaggc	gtgtcggcag	11040
219	gcttcggaaa	aattgtaggt	tgaaaagccg	aatatcgtcg	gaacgccgca	ggaaagcggt	11100
220	tcgatgatgt	tctgacaacc	cgaatcgacc	agactgccgc	cgacaaaagc	gacatcggcg	11160
221			ctcgcccata				11220
222	tegaceggea	aaccgtcgct	gcgccgctga	accttaaacc	cgaagcgttt	tgccgtttca	11280
223			gggatggcgc				11340
224			cagtttttcc				11400
225			coggootoog				11460
226	•		gatgtcgtat		-		11520
227			egeegeatee				11580
228			caggcggcgg		-	-	11640
229			aaacagcggc				11700
230			catcaaaatg				11760
231		-	gtcatacgga				11820
232			cgtcggggtc				11880
233			tatcaaggac				11940
±34			ggtaacggga				12000
235			ggcactteeg				12060
23h			tacatcataa				12120
.: 37			gttcagacgg				12180
238			acettgeacg				12240
239			gtgtgcgcgc				12300
240	tggcaggcag	acgttcggtc	gtgtageegt	ccaagaacgt	aatcgccgaa	gccatgcagg	12360

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/018,470A

TIME: 02/27/2003
TIME: 08:05:18

Input Set : N:\Crf4\02132003\J018470.raw
Output Set: N:\CRF4\02272003\J018470A.raw

#### Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the  $\langle 220 \rangle$  to  $\langle 223 \rangle$  fields of each sequence which presents at least one n or Xaa.

Seq#:6; N Pos. 6,8,38,39,96,102,341,390,410 Seq#:7; Xaa Pos. 3,13,114,137 Seq#:40; N Pos. 364,365,366,367,368,369,370,371,372,373,374,375,376,377,378 Seq#:40; N Pos. 379,380,381,382,383,384,385,386,387,388,389,390,391,392,393 Seq#:40; N Pos. 394,395,396,397,398,399,400,401,402,403,404,405,406,407,408 Seq#:40; N Pos. 409,410,411,412,413,414,415,416,417,418,419,420,421,422,423 Seq#:40; N Pos. 424,425,426,427,428,429,430,431,432,433,434,435,436,437,438 Seq#:40; N Pos. 439,440,441,442,443,444,445,446,447,448,449,450,451,452,453 Seq#:40; N Pos. 454,455,456,457,458,459,460,461,462,463,464,465,466,467,468 Seq#:40; N Pos. 469,470,471,472,473,474,475,476,477,478,479,480,481,482,483 Seq#:40; N Pos. 484,485,486,487,488,489,490,491,492,493,494,495,496,497,498 Seq#:40; N Pos. 499,500,501,502,503,504,505,506,895,982 Seq#:41; Xaa Pos. 122,123,124,125,126,127,128,129,130,131,132,133,134,135 Seq#:41; Xaa Pos. 136,137,138,139,140,141,142,143,144,145,146,147,148,149 Seq#:41; Xaa Pos. 150,151,152,153,154,155,156,157,158,159,160,161,162,163 Seq#:41; Xaa Pos. 164,165,166,167,168,169,170,171,172,173,174,175,176,177 Seq#:41; Xaa Pos. 178,179,180,181,182,183,184,185,186,187,188,189,190,191 Seq#:41; Xaa Pos. 192,193,194,195,196,197,198,199,200,201,202,332,361 Seq#:46; N Pos. 995,1082 Seq#:47; Xaa Pos. 332,361 Seq#:50; N Pos. 1216,1334,1347 Seq#:51; Xaa Pos. 149,158,167,190,198,289,313,328,331,339,343,347,406,445 Seq#:63; Xaa Pos. 116

VERIFICATION SUMMARY

DATE: 02/27/2003 TIME: ()8:05:18 PATENT APPLICATION: US/10/018,470A

Input Set : N:\Crf4\02132003\J018470.raw Output Set: N:\CRF4\02272003\J018470A.raw

L:38026 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:0 M:341 Repeated in SeqNo=6  $L:38056\ M:341\ W:$  (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:0 M:341 Repeated in SeqNo=7 L:39195 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:40 after pos.:360 M:341 Repeated in SeqNo=40 L:39234 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:41 after pos.:112 M:341 Repeated in SeqNo=41 L:39450 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:46 after pos.:960 M:341 Repeated in SeqNo=46 L:39507 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:47 after pos.:320M:341 Repeated in SeqNo=47 L:39628 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:50 after pos.:1200 M:341 Repeated in SeqNo=50

L:39712 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:51 after pos.:144

L:40420 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:63 after pos.:112

M:341 Repeated in SeqNo=51

US 1001847008P1



Creation date: 10-15-2003

Indexing Officer: ATEKLY - ALEM TEKLAY

Team: OIPEScanning Dossier: 10018470

Legal Date: 03-05-2003

No.	Doccode	Number of pages
1	M903	2

Total number of pages: 2

Remarks:

Order of re-scan issued on .....